Comparison of Time-Dependent Sequential Logit and Cox Proportional Hazards Models for Hurricane Evacuation with a Focus on the Use of Evolving Forecast Information

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Shuo Wang

December 2017

ABSTRACT

Hurricane is a natural disaster which could cause many deaths and considerable damage if improper emergency management was applied. Figuring out an efficient method to dynamically forecast the hurricane evacuation demand with high accuracy plays a crucial role in preparedness work of hurricane management. Recently, substantial studies and research exit on understanding hurricane evacuation behavior. However, in this thesis, some forecast covariates which were not mentioned before, are introduced into the prediction of hurricane evacuation rate. Moreover, two travel demand models are applied in this study: A Sequential Logit Model and a Cox proportional hazards model. These two models are used for estimating the probability of each household to evacuate in the specific time step. After applying the data from Hurricane Gustav (2008) in Louisiana, over 76% households' dynamic evacuation behavior are predicted correctly.

BIOGRAPHICAL SKETCH

Shuo Wang was born in Zhenjiang, Jiangsu, the Southeast part of China. She was graduated from the College of Transportation, Southeast University in Nanjing, with a bachelor degree in Transportation Engineering. With a keen interest towards transportation modeling and analysis, she decided to pursue a master's degree in the School of Civil and Environmental Engineering at Cornell University, with a focus on Transportation System Engineering.

Shuo tried different approaches to understand this dynamic field better and looking forward to making her contributions. She worked in Delaware Valley Regional Planning Commission during the summer of 2016, as a transportation modeling intern. Moreover, to enrich her professional experience, she took courses like Urban Transportation Network Design & Analysis, Micro econometrics of Discrete Choice, etc. at Cornell University. These courses and experience considerably broadened her horizon and helped her to further pursue a career in the field of Transportation.

This work is dedicated to

my parents, who always give me unconditional support and love,

and my advisors, for their generous patience and guidance.

ACKNOWLEDGMENTS

I would like to gratefully and sincerely thank Professor Linda K. Nozick, my special committee chair, for teaching me and giving me so many helpful advice. She is not only an expert in the transportation industry with profound knowledge and experience specifically in the development of mathematical models but also a pioneer in studying transportation and logistics systems. With her insightful suggestions and patient guidance, I found my interest in transportation modeling and analysis and research it. What is more, Prof. Nozick spent a lot of time reviewing the drafts, provided helpful revision advice, assisted with each step of my research. The thesis could not be done without her help.

I would like to express my sincere gratitude to Professor Huaizhu Oliver Gao for joining my special committee. From his class focusing on systems engineering, I developed an interest in using statistics and econometrics method to solve transportation problems. This knowledge helped me with my thesis topic and method selection.

I would also like to thank the Louisiana Transportation Research Center (LTRC) and The Public Policy Research Lab (PPRL), which provided the 2009 LSU Hurricane Evacuation Survey.

In the end, I would like to thank Tarun Rambha, the postdoc of Engineering School. He helped me a lot in the model choice and thesis writing. I am grateful to my parents and all my friends for supporting me financially and emotionally in completing the thesis.

TABLE OF CONTENTS

I. INTRODUCTION **5**

II. LITERATURE REVIEW **7**

2.1 FACTORS AFFECTING HURRICANE EVACUATION 7

2.1.1 Factors of the Household *7*

2.1.2 Factors of the Hurricane *8*

2.2 MODELS PROVIDED TO PREDICT EVACUATION RATE 11

2.2.1 The Probit Model *12*

2.2.2 The Logit Model *12*

2.2.3 Other Models *13*

III. DATA **15**

3.1 DATA COLLECTION 15

3.2 DATA DESCRIPTION 16

IV. METHODOLOGY **22**

4.1 SEQUENTIAL LOGIT MODEL 22

4.2 COX PROPORTIONAL HAZARDS MODEL 23

V. RESULTS OF MODEL ESTIMATION **24**

5.1 ESTIMATION RESULTS 24

5.2 PREDICTIVE POWER 27

VI. CONCLUSION **29**

REFERENCE: **31**

I. INTRODUCTION

Hurricane is a storm combined with strong winds and heavy rain and often occurs in the Atlantic or the northeastern Pacific Ocean. Hurricane can lead to a natural disaster which has a serious threat to people's lives and property. In history, records showed a great amount of hurricane happened in the United States. The Galveston Hurricane of 1900 was caused at least 8,000 people's death, which was called the deadliest hurricane. Twenty-eight years later, the Hurricane Okeechobee killed more than 2,500 casualties. In the 1893 season, two hurricanes led to the deaths of over 1,000 innocents. And in 2005, Hurricane Katrina killed about 1,500 people. All these horrible numbers and loss caused by hurricanes can be reduced or avoided if some relevant actions were taken in time.

Evacuation is one of the most practical solutions in the face of upcoming disasters. A proper evacuation plan can be developed to ensure the safety of all expected residents of the region threatened by hurricane disaster. Moreover, (Hasan *et al*., 2012) proves that the parameters of one evacuation choice model can also be used in other hurricanes. So, it is meaningful and reasonable for researchers to work on prediction of hurricane evacuation demand. A more geographically precise prediction of evacuation rates is necessary for the future. It will give a better idea of solid evacuation plans and accurate Hurricane consequences estimation. To improve the preparedness, there is a need to propose an efficient and precise predicted method of hurricane evacuation rate.

Moreover, including time as a factor in evacuation model is required. Some people did not evacuate in time because of the traffic jam. If we can forecast the evacuation period for each household or the evacuation rate in a particular time interval, it will be easier and more efficient to do the traffic-clearing work.

Section 2 summarizes the recent literature on hurricane evacuation. At the same time, it is divided into two parts. For the first part, a series of paper are provided to give a brief statement of main factors which affect the hurricane evacuation. And the other part is mainly about the researcher's dedicated work on finding the optimal model to predict the correlation between those factors and final decisions. In section 3, a detailed data collection and process for Hurricane Gustav are introduced, which is the dataset used the in the following sections to do the model estimation. Then, two models are proposed here in section 4 to predict the evacuation rates for the future hurricane. The first model is the Sequential Logit model which was presented in the paper (Fu and Wilmot, 2004). The second one is a Cox proportional-hazards Regression Model (Cox, 1992) which is usually used for survival analysis. In the end, the final fitted models are represented. And a comparison of estimated parameters and the prediction performance of these two models are shown then. As a result, both models perform well in predictive power. Over 76% households' dynamic evacuation behavior are predicted correctly.

II. LITERATURE REVIEW

2.1 FACTORS AFFECTING HURRICANE EVACUATION

Throughout the past few decades, much progress has been made in the field of the hurricane evacuation. Researchers have done plenty of works on the evacuation behavior and tried to find all key factors which affect evacuation decisions.

Baker proposed five influential factors in hurricane evacuation (Baker, 1991): risk level of the area, actions by public authorities, housing, a prior perception of personal risk and storm-specific threat factor. These five factors were acknowledged and identified by most researchers. However, it's a complex problem that why some of the people evacuated from a disaster, while others choose to stay. This is an interdisciplinary topic which covers behavioral sciences, psychology, and engineering. So, the factors which have influences on the evacuation behavior include multiple fields. A more detailed review about the intense literature focus on factor-selection is listed below.

2.1.1 Factors of the Household

After analyzing the survey data by using logistic regression model, (Riad *et al.*, 1999) proposed three main social psychological aspects: risk perception, social influences and access to resources. They thought the characteristics and these three processes have the essential impacts on the evacuation behavior simultaneously. Three years later, (Bateman and Edwards, 2002) explained the influence of gendered variations on evacuation decisions in their paper. They summarized the results of the cross-validation survey of households affected by Hurricane Bonnie, on August 25, 1998. So, females are more likely to evacuate when in danger of the hurricane. Because, they play the caregiving roles, and have more sharp risk-perception than man. Additionally, (Lazo *et al.*, 2010) concentrated on analyzing the evacuation behavior concerning hurricane characteristics and warning information. Their research showed that people were willing to pay for the improvement of the forecast accuracy, especially in the field of landfall timing and location, storm surge and winds speed. Similarly, (Reininger *et al.*, 2013) tried to find out people's reaction to the hurricane evacuation orders. Findings show that most residents intend to follow the mandatory evacuation orders for all income levels. However, the variation of comply or not may exist if those residents have differences in age, gender, distance from the shoreline.

In recent years, some new factors were presented with the change of people's habit and life. (Widener *et al.*, 2012) employed an ABM with geographic data and assessed how social network structures influence the rate of participation in hurricane evacuation. It can increase the number of evacuating residents to some extent. (Huang *et al.*, 2012) mentioned in their study that departure timing is an important factor which cannot be neglected.

2.1.2 Factors of the Hurricane

Similarly, (Whitehead *et al., 2000)* mentioned the effect of storm intensity which plays a major role in the hurricane evacuation. Florida experienced a bad hurricane season in 2014. (Smith and Mccarty, 2009) collected the demographic data about the households influenced by four hurricanes happened in 2014 summer, and found that the characteristics of hurricane and house had the greatest impact on evacuation behavior. Apart from that, a conditional mutual information maximization method was proposed to deal with the factors' selection. In this research, "distance" is found to be one of the factors with highest predictive power. (*Demiroluk et al.*, 2016)

In summary, some useful overview paper also gives an outline of some important factors that affect evacuation behavior, such as (Yazici and Ozbay, 2008) and (Dash and Gladwin, 2007).

Table 1 summarizes the factors stated to affect evacuation behavior and demand, including the studies that cited these factors.

Significant Factor	Related Study
Location	Kecheng Xu et al., 2016; Sami Demiroluk et al., 2016; Petrolia & Bhattacharjee, 2010; Reininger et al., 2013; Gudishala and Wilmot, 2012; Hasan, et al., 2013; Fu and Wilmot, 2004; Yin et al., 2014; Lim et al., 2016; Wilmot and Mei, 2004; Whitehead, 2005; Widener et al., 2012; Russo and Chilà, 2014
Housing Type	Kecheng Xu et al., 2016; Yin et al., 2014; Hasan, et al., 2013; Petrolia & Bhattacharjee, 2010; Wilmot and Mei, 2004; Fu and Wilmot, 2004; Lazo et al., 2010; Whitehead, 2005; Lim et al., 2016; Lim et al., 2016; Yin et al., 2014
Home ownership	Sami Demiroluk et al., 2016; Lim et al., 2016; Lazo et al., 2010
Gender	Kecheng Xu et al., 2016; Whitehead, 2005; Petrolia & Bhattacharjee, 2010; Riad, et al., 1999; Bateman and Edwards, 2002; Lazo et al., 2010; Reininger et al., 2013; Ng et al., 2014; Sami Demiroluk et al., 2016; Lim et al., 2016; Russo and Chilà, 2014
Race/ Ethnicity/ Cultural issues	Kecheng Xu et al., 2016; Whitehead, 2005; Riad, et al., 1999; Lazo et al., 2010; Petrolia & Bhattacharjee, 2010; Hasan, et al., 2013; Ng et al., 2014; Sami Demiroluk et al., 2016; Reininger et al., 2013
Education	Kecheng Xu et al., 2016; Petrolia & Bhattacharjee, 2010; Reininger et al., 2013; Sami Demiroluk et al., 2016; Hasan, et al., 2013; Lim et al., 2016; Lazo et al., 2010; Yin et al., 2014; Whitehead, 2005
Job	Kecheng Xu et al., 2016; Lazo et al., 2010; Russo and Chilà, 2014; Russo and Chilà, 2014; Yin et al., 2014

Family size	Wilmot and Mei, 2004; Petrolia & Bhattacharjee, 2010; Lazo et al., 2010; Sami Demiroluk et al., 2016; Ng et al., 2014
Presence of Child	Kecheng Xu et al., 2016; Yin et al., 2014; Lim et al., 2016; Ng et al., 2014; Sami Demiroluk et al., 2016; Hasan, et al., 2013
Presence of Elderly	Ng et al., 2014
Pet ownership	Whitehead, 2005; Lazo et al., 2010; Petrolia & Bhattacharjee, 2010; Sami Demiroluk et al., 2016; Yin et al., 2014
Age	Kecheng Xu et al., 2016; Petrolia & Bhattacharjee, 2010; Reininger et al., 2013; Russo and Chilà, 2014; Wilmot and Mei, 2004
Income	Kecheng Xu et al., 2016; Hasan, et al., 2013; Whitehead, 2005; Petrolia & Bhattacharjee, 2010; Gudishala and Wilmot, 2012; Lazo et al., 2010; Yin et al., 2014; Ng et al., 2014; Reininger et al., 2013
Previous Experience	Riad, et al., 1999; Petrolia & Bhattacharjee, 2010; Ng et al., 2014; Widener et al., 2012; Sami Demiroluk et al., 2016; Lazo et al., 2010
Length of residence	Ng et al., 2014; Sami Demiroluk et al., 2016; Lim et al., 2016; Lazo et al., 2010
Storm properties	Petrolia & Bhattacharjee, 2010; Gudishala and Wilmot, 2012
Flood risk & Wind risk	Whitehead, 2005; Fu and Wilmot, 2004; Gudishala and Wilmot, 2012; Ng et al., 2014; Lim et al., 2016; Lazo et al., 2010
Dissemination of orders	Kecheng Xu et al., 2016; Whitehead, 2005; Petrolia & Bhattacharjee, 2010; Hasan, et al., 2013; Fu and Wilmot,

	2004; Gudishala and Wilmot, 2012; Sami Demiroluk et al., 2016; Yin et al., 2014; Wilmot and Mei, 2004; Lim et al., 2016
Accuracy of hurricane forecast	Lazo et al., 2010
Strong social network	Ng et al., 2014; Riad, et al., 1999
Vehicles owned	Russo and Chilà, 2014; Gudishala and Wilmot, 2012; Sami Demiroluk et al., 2016; Widener et al., 2012; Lazo et al., 2010
With driver license or not	Russo and Chilà, 2014
Disabled or not	Lazo et al., 2010; Sami Demiroluk et al., 2016; Petrolia & Bhattacharjee, 2010
Have specific evacuation destination or not	Lazo et al., 2010; Ng et al., 2014; Petrolia & Bhattacharjee, 2010
Confidence in being rescued	Petrolia & Bhattacharjee, 2010
Time of day	Gudishala and Wilmot, 2012
Acculturation	Reininger et al., 2013
Time to destination	Hasan, et al., 2013; Ng et al., 2014
Burglary chance	Ng et al., 2014
Preparedness	Ng et al., 2014; Sami Demiroluk et al., 2016

2.2 MODELS PROVIDED TO PREDICT EVACUATION RATE

The previous review indicates some significant factors which affect decisions of evacuation or not, even the timing of departure. However, a reasonable model needs to be found to predict the evacuation rate. Generally, three types of model were proposed in the previous studies, probit model, logit model, and neural network model, etc.

2.2.1 The Probit Model

(Whitehead, 2005) proposed a probit model to predict the evacuation rates by using data collected in and before 1999 hurricane season in Carolina. It is proved that stated behavior data for hurricane evacuations are predictive valid. Some jointly estimated revealed and stated behavior models were used to test and compare these two surveys. As a result, the forecast error is small which indicates stated behavior data has the predictive validity and the probit model is feasible in evacuation choice prediction problem. A new ordered probit model is proposed by Kecheng Xu to help with evacuation rate prediction for any future hurricane under different order type. The most important feature of this model is that all covariates can be easily approached from the census prone. Cross validation analysis also showed the predicted accuracy of this ordered probit model is acceptable high. (Xu *et al*., 2016)

2.2.2 The Logit Model

(Petrolia & Bhattacharjee, 2010) proposed to use multinomial logit model analyzing data from hypothetical storm forecast scenarios. This study focuses on the differences between people who are in the group of non-evacuate instead of the differences between evacuate and non-evacuate. All respondents have three choices: wait, don't evacuate and don't know. The results of this study can help us encourage people to make decisions quickly when facing a hurricane. Similarly, (Lim et al., 2016) formulated a model allows a choice among three alternatives of full, partial, and no evacuation. A time dependent sequential logit was formulated by (Fu and Wilmot, 2004), but it has some restrictive assumptions. A time-dependent nested logit model was proposed by (Gudishala and Wilmot, 2012), which relaxes those assumptions. (Hsu and Peeta, 2013) proposed an aggregate behavior model for evacuation decision and evacuation route choice to support information-based control for the real-time stage-based routing of individuals in the affected areas. By applying a mixed logit structure, this model accounts for the heterogeneity across the evacuees.

2.2.3 Other Models

(Wilmot and Mei, 2004) compared the performance of conventional participation rate, logistic regression and neural network models to get an idea about which model can provide a more accurate result. The research indicates that logistic regression and neural network models have higher accuracy in prediction than the participation rate model.

(Fu and Wilmot, 2006) proposed two dynamic travel demand models for hurricane evacuation based on survival analysis: A Cox proportional hazards model and a piecewise exponential model. And the data from Hurricane Andrew in southeast Louisiana was used to examine these two model. A piecewise exponential model performed better with less limitation in the end.

The choice of departure time during the hurricane is a complicated dynamic process and relies on many features. For example, different households with unique characteristics which lead them to leave earlier or later. Moreover, the timing of evacuation issued also influences the period they choose to depart. (Hasan, et al., 2013) developed a random-parameter hazard-based model to predict hurricane evacuation timing by individual households.

Similarly, (Ukkusuri1 et al., 2016) presented A-RESCUE, a high fidelity multi-agent simulation model that integrates household-level activity behavior with a network-level traffic assignment to evaluate a broad range of evacuation strategies.

It is worth to mention that (Russo and Chilà, 2014) proposed a general framework about what to do with the evacuation conditions in dangerous. Most of the papers are only concentrate on one particular formal transportation decision. However, this article provides a general structure of how to simulate all kinds of transportation decisions. A demand model specified and calibrated with influences on transport demand according to behavioral analysis.

III. DATA

3.1 DATA COLLECTION

A survey to get evacuation metadata about Hurricane Gustav was conducted by Public Policy Research Lab with the sponsorship from the Louisiana Transportation Research Center. This survey was completed from July 23, 2009 to October 20, 2009 in two stages. Households in the first stage were enlisted through telephone, which was from ten parishes near New Orleans. In the second stage, the questionnaires were mailed to all households pass the first stage which was willing to help with the survey. Also, all these households were required to have experienced Hurricane Gustav when it made landfall in September 2008. The parishes of Tangipahoa, St. John the Baptist, Plaquemines, Jefferson, Orleans, St. Tammany, Lafourche, St. Charles, Terrebonne, and St. Bernard, were included in the population. Time-dependent, hurricane evacuation behavior data were collected then.

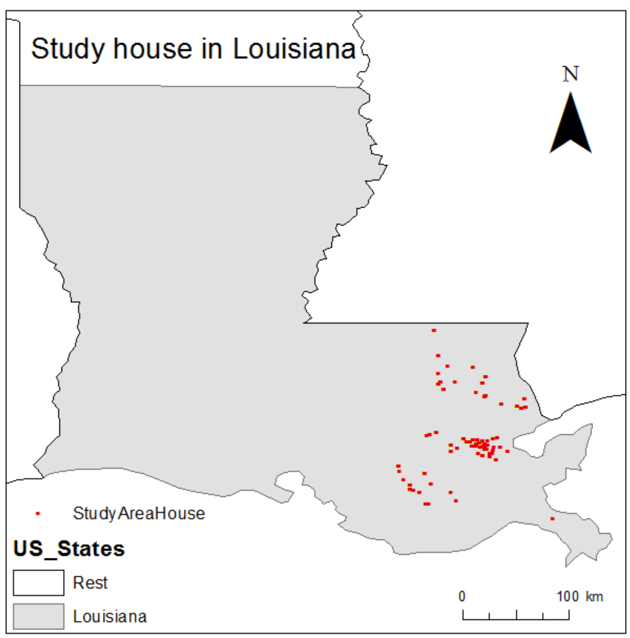

Fig. 1 Visualization of location of observation households

Apart from that, some dynamic data was collected with Hurricane Gustav's storm-related information by retrieving from the archives of the National Hurricane Center.

3.2 DATA DESCRIPTION

Demographic characteristics measured included household size, the highest level of schooling, household income and ethnicity (Asian, Black African/American, Indian(American), Mixed race, White, Other). Also, to measure more characteristics associated with natural hazard-related preventive behavior, some other features are included, such as:

(1) Whether the respondent residents living in the flood zone;
(2) Type of house the household was living when Gustav made landfall (Permanent house, Mobile house, Apartment/Condo, Other);
(3) Number of vehicles owned;
(4) Job required you to stay in the area during evacuation;
(5) Household size
(6) Number of people who are 17 or younger living in household;
(7) Have pets or not;
(8) The number of years resided at the current residence.

For those observations including variables with no response value, we discarded them to ensure the reliability of model estimation. After cleaning the data, 244 households retained. Of these households, 173 of them chose to evacuate during the Hurricane Gustav. Reasons for respondents to evacuate or stay were asked in the questionnaires too. There are some reasons listed for respondents to choose. **Fig.2 and Fig.3** show the percentage of each reason for evacuating and staying, which can help us to get a general sense of the phenomenon that why some of the residents choose to evacuate, however, others choose to stay. Apparently, evacuation orders from emergency and elected officials, advice from weather service and concerned flooding would flood home or cut off roads are the top three reasons for evacuating. And for people who choose to stay, storm properties and house characteristics are the most important factors. Naturally, in the next section of hurricane evacuation prediction, all these significant reasons will be taken into consideration.

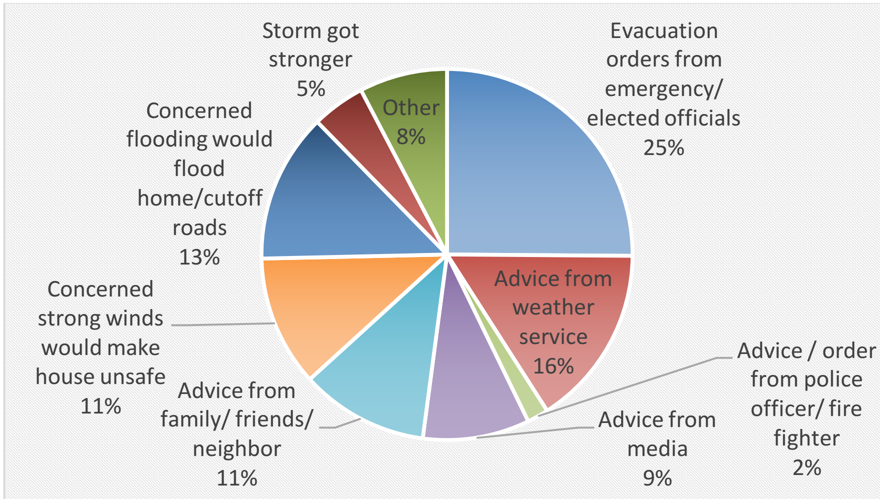

Fig. 2 Percentage of different reasons for evacuating

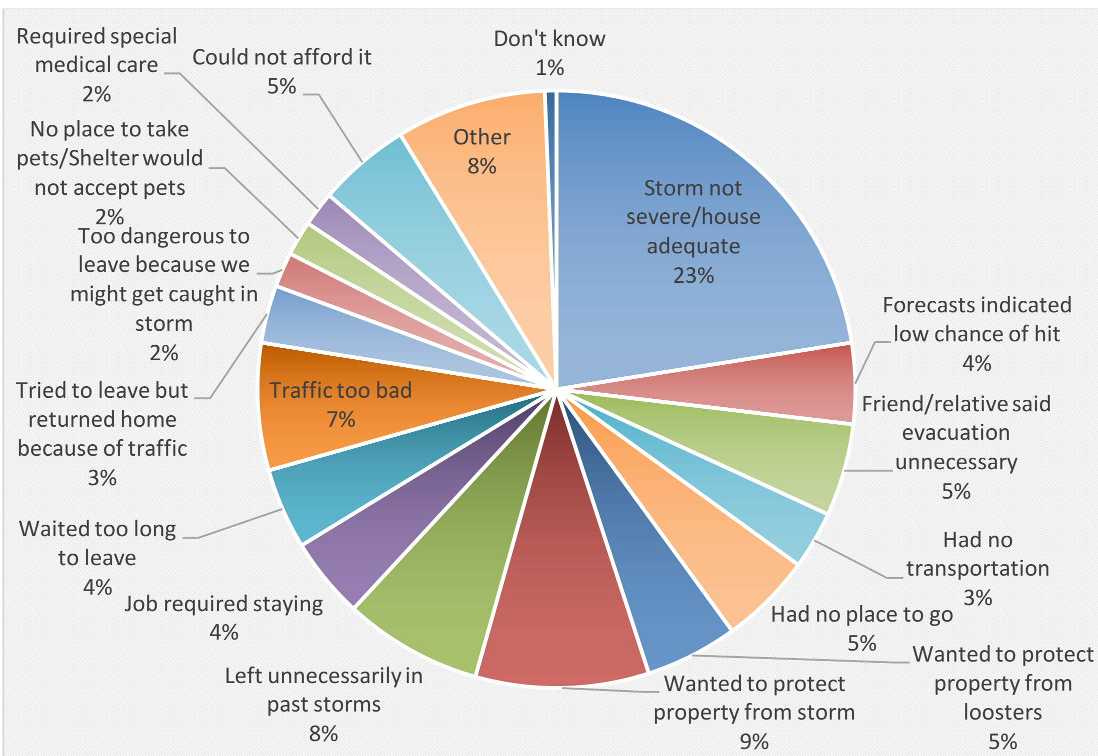

Fig. 3 Percentage of different reasons for staying

Table 2 and **Table 3** summarize the descriptive statistics for interval and categorical variables used in the following analysis. According to the previous research and literature, 11 variables were extracted which perform better in predicting evacuation rates. For the household income, due to all the observations are reported as a categorical variable, that is, a range criterion was used, we took the midpoint of the range to do the analysis.

Table 2 Summary statistics for interval variables

Variable definition	Variable	Mean	SD
No. of vehicles owned	NO_VEH_OWN	2.05	1.33
Household size	HHSIZE	2.77	1.25
No. of people who are 17 or younger living in household	NO_YOU_17	0.54	0.92
Number of years resided at current resident	LEN_RES_YRS	17.43	14.66
Total household income per year($)	HHINC	75.27664	46.91787

Household income: 1 = Less than $15k, 2 = $15k to 25k, 3 = $25k to 40k, 4 = $40k to 80k, 5 = $80k to 120k, 6 = $120k to 150k, 7 = Over $150k.

Table 3 Number (%) of response for each categorical variable

Variable definition	Variable	Category	*f*	%
Location	FL_ZONE	Respondent resided in the flood zone	8	3.28%
	-	Respondent did not reside in the flood zone	236	96.72%
Housing type	Permanent	Permanent house	223	91.39%
	-	Mobil home/ Apartment/ Condo and Other	21	8.61%

Job	JOB_01	Job requires you to stay in the area during evacuation	26	10.66%
	-	Job does not require you to stay in the area during evacuation	218	89.34%
Pet	PETS_01	Have pets	136	55.74%
	-	No pets	108	44.26%
Education	BachelorAndHigher	Bachelor's degree or higher	171	70.08%
	-	Other	73	29.92%
Ethnicity	White	White	209	85.66%
	-	Other	35	15.98%

Then, those dynamic data collected by the National Hurricane Center takes the following variables at every time interval into consideration: hurricane category, evacuation orders issued or not, time-dependent distance, time-of-day, storm surge, forecast distance, winds speed and predicted the probability of storm surge for the survey area.

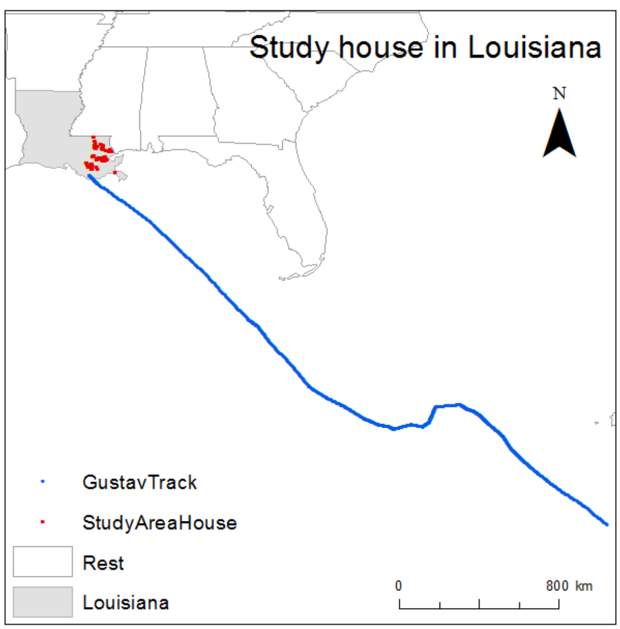

Fig. 4 Hurricane Gustav track from National Hurricane Center

Data collected above was rearranged for the dynamic travel demand model. That is, in which time step that household would choose to leave. Each row of observations from a single household was expanded into multiple rows. In the expanded data, each row represented a time step of 6 hours, and the dynamic variables (order, hurricane category, time of day, time-dependent distance, storm surge etc.) varied between these time periods. The number of expanded rows was determined by in which period this household decides to evacuate. For example, if the household left in time step 18, the observation for this household would be expanded into 18 rows.

All those time-dependent variables considered in the model are shown below.

1. *Distance:* The distance from the center of the hurricane to the geographical location of the household. However, this variable doesn't affect people's action of evacuating linearly. For example, if a household resides far away from the coastal line, decreasing a unit of distance will not affect people there a lot. On the contrary, if the people live near the water, it will be much more dangerous for him/her if the distance changes a unit. Apart from that, once the distance of a hurricane is less than a threshold, it will be too dangerous to leave their house. Here we set d_{min} as 94mi. Then, a transformed value of distance was used as below:

$$f(t) = \begin{cases} 0 & \text{if } d(t) < d_{min} + 1 \\ \ln[d(t) - d_{min}] & \text{otherwise} \end{cases}$$

2. *Hurricane Category:* Five categories of hurricane corresponding to the Saffir-Simpson scale. A larger value represents a more severe hurricane.
3. *Evacuation Order:* An evacuation order refers to the action taken by public officials specifying the type and timing of an evacuation order issued. It was entered as a dummy variable. A mandatory or voluntary order was represented by 1, and no evacuation order was represented by 0.
4. *Time of Day:* Three dummy variables, TOD1, TOD2, and TOD3 were represented. If the time-of-day was between 12 am and 6 am then the TOD1

was coded as 1 and zero otherwise. If the time-of-day was between 6 am and 12 pm then TOD2 was coded as 1 and 0 otherwise. TOD3, represented the time between 12 pm to 6 pm and was coded as 1 if the time of day fell in that category and 0 otherwise. The time between 6 pm and 12 am was used as the base and was represented in the data with zeros on TOD1, TOD2, and TOD3.

5. *Storm surge:* A dummy variable. If the value of storm surge is 1, the hurricane results in flooding depth greater than 10 ft. Because the 10 ft is a mean of home sites higher than ground level.
6. *ForecastDist:* Variable Forecast records the distance from each observation household to forecast hurricane landfall location over time in kilometers.
7. *Winds:* Winds1, winds2 and winds3 represent the probability of hurricane force surface winds 1-minute average greater than 39, 58 and 74mph respectively. Category 0 - 10 is set based on the legend as following. If the probability is 5%-10%, it is set as category 1 in the spreadsheets; if the probability is 10%-20%, it is set as category 2 and so on.
8. *Psurge:* Psurge is a categorical variable which represents the probability of storm surge. Category 0 - 10 is set based on the same legend as variable winds.

IV. METHODOLOGY

Two models are proposed to predict the dynamic evacuation rates for a future hurricane. The first model is the Time-dependent Sequential Logit model which was presented in paper Fu, H., & Wilmot, C. (2004). The second one is a Cox Proportional Hazards Model which is traditionally used for survival analysis.

4.1 SEQUENTIAL LOGIT MODEL

When evacuating in the face of an oncoming hurricane, people usually experience a mental activity about which period to leave in as follows. A household needs to decide either to evacuate or stay in each period i, and these decisions occur sequentially over time until the household evacuates in one of these periods or stays in the end. Moreover, sequential choice happens when next choices can be realized only by first experiencing earlier choices. (Fu and Wilmot, 2006).

Naturally, the probability that a household will evacuate at period i can be expressed in the form of a regular binary logit model for each period i. After applying the utility theory, a specific utility is assigned to each household. The utility of household n choosing to stay in period i is:

$$\mathrm{U}_{nsi} = \beta x_{nsi} + \varepsilon_{nsi}$$

Where β is a vector of parameters, ε_{nei} is the unobserved error term which is independently and identically Gumbel distributed.

Similarly, the utility of household n choosing to evacuate in period i is:

$$\mathrm{U}_{nei} = \beta x_{nei} + \varepsilon_{nei}$$

Then, assuming the decision of period i is not affected by other periods, the probability that household n will choose to evacuate in period i can be considered as a binary logit model:

$$P_{nei} = \frac{e^{\beta x_{nei}}}{e^{\beta x_{nsi}} + e^{\beta x_{nei}}}$$

However, the decision to evacuate is a joint event which in the condition that the household chose to stay from period 1 to period $(i - 1)$. So, the probability that a household will evacuate at period i:

$$p(i) = p(i)_{evac} \prod_{j=1}^{i-1} [1 - P(j)_{evac}]$$

And the parameter β can be estimated from a sample of N household by using the following likelihood function:

$$L(\beta) = \prod_{n=1}^{N} P_{nei} \prod_{j=1}^{i-1} [1 - P_{nej}]$$

4.2 COX PROPORTIONAL HAZARDS MODEL

Survival analysis is generally defined as a set of methods for analyzing data where the outcome variable is the time until the occurrence of an event of interest. (Kleinbaum D.G., 1996). In this thesis, the event is considered as the hurricane evacuation.

One of the most significant advantages of survival analysis over ordinary regression models is to use information from censored data. Survival methods can estimate parameters by incorporating information from both censored and uncensored observations. And the data consist of event condition and time to such event.

There are two essential functions in survival analysis which are the survival function and the hazard function. The survival function provides the probability of choosing to stay (not evacuate) up to that time. The hazard function gives the potential that people will evacuate, per time unit, given that they have stayed up to the specified time.

In this thesis, a popular regression model, Cox proportional hazards regression model, is taken into consideration. The detailed explanation of applying Cox proportional hazard model in predicting the evacuation rate for Hurricane is shown below.

For household $i = 1, \dots, n$, let E_i denote the evacuation time. C_i denote the censoring time. Here we censored 19 times which were the 19 time steps. $N_i(E)$ represent a counting process such that $N_i(E) = I\,(E_i \geq t)$, where $I\,(u)$ is the indicator function taking value 1 if event u occurs and 0 otherwise. A household is at risk until they evacuate or are censored. $Y_i(E) = I\{\min\,(E_i, C_i) < t\}$. Let X_i denote a vector of covariates. The evacuation time E_i is not available for all subjects, but instead min (E_i, C_i) and $\delta_i = I\,(E_i \leq C_i)$ are observed. The hazard of evacuation $h(t|X)$ is related to the covariates by

$$h(t|X) = h_0(t)\exp\,(\sum \beta X_i)$$

Where $h_0(t)$ is an unspecified baseline hazard function for the reference subject with all covariates equal to 0.

Moreover, a useful and easy to interpret information was provided regarding the relationship of the hazard function to predictors, which is the hazard ratio. The effect of one unit increase in X, is measured by hazard ratio $\frac{h(t|X=x+1)}{h(t|X=x)} = \exp\,(\beta)$.

V. RESULTS OF MODEL ESTIMATION

5.1 ESTIMATION RESULTS

A backward stepwise selection was used to choose the optimal model. Outcomes of the optimal sequential logit model and the Cox proportional hazards model are presented below:

Table 4 Summary Results of Model Estimation

	Sequential Logit Model			Cox Model		
Covariate	β	se(β)	p-Value	β	se(β)	p-Value
Time-dependent-distance	-0.62	0.08	5.15e-15	-2.31	0.30	2.12e-14
Hurricane category	0.18	0.09	0.04753	4.06	0.62	4.33e-11
TOD1	0.91	0.32	0.00485	1.30	0.30	1.60e-05
TOD2	1.63	0.20	1.07e-07	1.68	0.24	9.48e-13
TOD3	0.63	0.34	0.06072	-	-	-
Storm surge	0.67	0.39	0.08304	0.66	0.35	0.0568
ForecastDist	-0.01	0.00	3.27e-05	-0.01	0.00	4.21e-06
Winds1	0.32	0.04	4.99e-14	-	-	-
NO_VEH_OWN	-0.16	0.07	0.02397	-0.14	0.06	0.0303
LEN_RES_YRS	-0.02	0.01	0.00805	-0.01	0.01	0.0073
Ethnicity	-0.15	0.08	0.06669	-0.13	0.08	0.0768

All covariates have the same sign for both models and they all found to be significant at the 95% significant level in the sequential logit model. Two new forecast factors are included: the distance from each observation household to forecast hurricane landfall location, the probability of hurricane force surface winds 1-minute average greater than 39mph. Some covariates such as TOD3 and Winds1 are not significant in the Cox proportional hazards model, so they were not included in the right part. Moreover, the influence of evacuation order is also taken into consideration when estimating parameters for time-dependent-distance and hurricane category in cox model.

In general, all the covariates have the correct sign and the values are acceptable. For the signs of covariates Time-of-Day, they are all positive, which indicates that most households are willing to leave in daytime. Among all these three covariates, TOD2 has a larger coefficient in both models, that is, people are more likely to leave in the morning from 6:00 am to 12:00 pm.

About the sign of time-dependent-distance, it is negative as expected. It would be dangerous as the storm became closer. Naturally, people are more likely to leave as the distance decreases due to the higher risk degree. Similarly, hurricane category is positive in both models. That is, with increasing level of hurricane category, people have a higher probability to leave.

Among the rest covariates, storm surge has the largest parameter value. People living in the higher potential storm surge area are more likely to evacuate, which is mathematically more than 1.9 times greater in both models.

One of the two forecast covariates, predicted wind speed, has the positive sign. The likelihood of evacuation is 1.38 times higher if a probability of hurricane force surface winds 1-minute average greater 39mph increases.

Then, people are less likely to leave if the distance from each observation household to forecast hurricane landfall location increases. It makes sense that if the forecast landfall location of the hurricane is closer, the probability of being attacked will be higher. And evacuation becomes more necessary. And as expected, people are less likely to leave if they resident in this area for a long time. Also, the negative sign of the parameter associated with number of vehicles owned indicates that the more valuable of their home, the less likely they will leave.

An interesting result is also showed in the **table 4** that the ethnicity does effect the evacuation behavior in Hurricane Gustav. White people are less likely than other people to leave when facing the hurricane. However, this result needs to be examined and proved by more survey data of other hurricanes to reduce the errors caused by sampling.

Parameters of the logit model are interpreted as the effect on the likelihood of evacuating. Below is a summary of the quantitative results of the effect on increasing the likelihood of evacuation.

Table 5 Odds & Hazards Ratio

Variable	Odds Ratio	Hazards Ratio
Time-dependent-distance	0.54	0.10
Hurricane category	1.20	57.97
TOD1	2.48	3.67
TOD2	5.10	5.37
TOD3	1.88	-
Storm surge	1.95	1.93
ForecastDist	0.99	0.99
Winds1	1.38	-
NO_VEH_OWN	0.85	0.87
LEN_RES_YRS	0.98	0.99
Ethnicity	0.86	0.88

5.2 PREDICTIVE POWER

In this part, we randomly selected 2297 observations (62.13% in total observations) from the dataset as the training data. The remaining 1400 observations are treated as a testing dataset for the model validation.

After estimating the sequential logit model and Cox model respectively by training data, a prediction for hurricane evacuation choice is made by using the estimated parameters. Four error metrics of both models for individual household level prediction (Testing dataset) is shown in **table 6** and **table 7**.

Table 6 Four error metrics for the Sequential Logit prediction

	STAY	EVACUATE
PREDICTED TO STAY	1013	15
PREDICTED TO EVACUATE	354	53

Table 7 Four error metrics for the Cox prediction

	STAY	EVACUATE
PREDICTED TO STAY	1019	15
PREDICTED TO EVACUATE	346	53

Concretely, Cox model's corrected prediction to stay is a bit higher than sequential logit model, while the same in predicting evacuation. And the percentages of corrected predictions of stay and evacuate are both more than 76%. To conclude, both models have a good predictive power in a hurricane evacuation choice problem.

VI. CONCLUSION

In this thesis, some forecast covariates are introduced into the prediction of hurricane evacuation rate, for example, the distance from each observation household to forecast hurricane landfall location and the probability of hurricane force surface winds.

Moreover, two travel demand models are applied by using data collected from Hurricane Gustav in this study: A Sequential Logit Model and a Cox proportional hazards model which are based on survival analysis. From the result of model estimation, covariates time-dependent-distance, hurricane category, time-of-day, storm surge, forecast landfall location, number of vehicles and length of resident years are all found to be significant at the 95% significant level in both models.

To conclude, people are less likely to leave if the distance from each observation household to forecast hurricane landfall location are far away. And as expected, people are less likely to leave if they resident in this area for a long time. Also, the negative sign of the parameter associated with number of vehicles owned indicates that the more valuable of their home, the lower probability they will leave.

People living in the higher potential storm surge area are more likely to evacuate, which is mathematically more than 1.9 times higher in both models. A similar influence caused by dummy variables TOD1, TOD2 and TOD3 which indicates that most households are willing to leave in daytime, especially in the morning. One of the forecast covariates predicted wind speed also has positive effects to evacuation. And for the same household, as the degree of risk becoming higher, they will less likely to leave their home.

In the end, estimation results for these two models are applied to predict each household's choice in the specific time step. As a result, over 76% households' dynamic evacuation behavior are predicted correctly for both models which are equally accurate in predicting the dynamic evacuation demand of Hurricane Gustav.

For future research, a nested logit model is an appropriate idea for the evacuation demand prediction which was tried during this study. Because when a household decided to stay, it is reasonable to assume that the conditions of next time periods have been taken into consideration. So, the utility of each household in a particulartime step should include future time steps' utilities in.

REFERENCE:

Baker, E. J. (1991). Hurricane evacuation behavior. *International Journal of Mass Emergencies and Disasters*, *9*(2), 287-310.

Riad, J. K., Norris, F. H., & Ruback, R. B. (1999). Predicting evacuation in two major disasters: risk perception, social influence, and access to Resources1. *Journal of Applied Social Psychology*, *29*(5), 918-934.

Whitehead, J., Edwards, B., Van Willigen, M., Maiolo, J., & Wilson, K. (2000). Hurricane evacuation behavior: a preliminary comparison of Bonnie, Dennis, and Floyd. *the Aftermath of Hurricane Floyd: Recovery in the Coastal Plain, Greensville, NC*.

Bateman, J. M., & Edwards, B. (2002). Gender and evacuation: A closer look at why women are more likely to evacuate for hurricanes. *Natural Hazards Review*, *3*(3), 107-117.

Smith, S. K., & McCarty, C. (2009). Fleeing the storm (s): An examination of evacuation behavior during Florida's 2004 hurricane season. *Demography*, *46*(1), 127-145.

Lazo, J. K., Waldman, D. M., Morrow, B. H., & Thacher, J. A. (2010). Household evacuation decision making and the benefits of improved hurricane forecasting: Developing a framework for assessment. *Weather and Forecasting*, *25*(1), 207-219.

Widener, M. J., Horner, M. W., & Metcalf, S. S. (2013). Simulating the effects of social networks on a population's hurricane evacuation participation. *Journal of geographical systems*, *15*(2), 193-209.

Huang, S. K., Lindell, M. K., Prater, C. S., Wu, H. C., & Siebeneck, L. K. (2012). Household evacuation decision making in response to Hurricane Ike. *Natural Hazards Review*, *13*(4), 283-296.

Reininger, B. M., Raja, S. A., Carrasco, A. S., Chen, Z., Adams, B., McCormick, J., & Rahbar, M. H. (2013). Intention to comply with mandatory hurricane evacuation orders among persons living along a coastal area. *Disaster medicine and public health preparedness*, *7*(01), 46-54.

Murray-Tuite, P., & Wolshon, B. (2013). Evacuation transportation modeling: An overview of research, development, and practice. *Transportation Research Part C: Emerging Technologies*, *27*, 25-45.

Demiroluk, S., Yazici, M. A., Ozbay, K., & Carnegie, J. A. (2016). Feature Selection for Ranking of Most Influential Variables for Evacuation Behavior Modeling Across Disasters. *Transportation Research Record: Journal of the Transportation Research Board*, (2599), 24-32.

Yazici, M. A., & Ozbay, K. (2008). Evacuation modelling in the United States: Does the demand model choice matter?. *Transport Reviews*, *28*(6), 757-779.

Dash, N., & Gladwin, H. (2007). Evacuation decision making and behavioral responses: Individual and household. *Natural Hazards Review*, *8*(3), 69-77.

Wilmot, C. G., & Mei, B. (2004). Comparison of alternative trip generation models for hurricane evacuation. *Natural hazards review*, *5*(4), 170-178.

Whitehead, J. C. (2005). Environmental risk and averting behavior: Predictive validity of jointly estimated revealed and stated behavior data. *Environmental and Resource Economics*, *32*(3), 301-316.

Fu, H., & Wilmot, C. (2006). Survival analysis-based dynamic travel demand models for hurricane evacuation. *Transportation Research Record: Journal of the Transportation Research Board*, (1964), 211-218.

Petrolia, D. R., & Bhattacharjee, S. (2010). Why don't coastal residents choose to evacuate for hurricanes?. *Coastal Management*, *38*(2), 97-112.

Hasan, S., Mesa-Arango, R., Ukkusuri, S., & Murray-Tuite, P. (2011). Transferability of hurricane evacuation choice model: joint model estimation combining multiple data sources. *Journal of transportation engineering*, *138*(5), 548-556.

Gudishala, R., & Wilmot, C. (2012). Comparison of time-dependent sequential logit and nested logit for modeling hurricane evacuation demand. *Transportation Research Record: Journal of the Transportation Research Board*, (2312), 134-140.

Hsu, Y. T., & Peeta, S. (2013). An aggregate approach to model evacuee behavior for no-notice evacuation operations. *Transportation*, *40*(3), 671-696.

Hasan, S., Mesa-Arango, R., & Ukkusuri, S. (2013). A random-parameter hazard-based model to understand household evacuation timing behavior. *Transportation research part C: emerging technologies*, *27*, 108-116.

Yin, W., Murray-Tuite, P., Ukkusuri, S. V., & Gladwin, H. (2014). An agent-based modeling system for travel demand simulation for hurricane evacuation. *Transportation research part C: emerging technologies*, *42*, 44-59.

Russo, F., & Chilà, G. (2014). Integrated travel demand models for evacuations: a bridge between social science and engineering. *International journal of safety and security engineering*, *4*(1), 19-37.

Xu, K., Davidson, R. A., Nozick, L. K., Wachtendorf, T., & DeYoung, S. E. (2016). Hurricane evacuation demand models with a focus on use for prediction in future events. *Transportation Research Part A: Policy and Practice*, *87*, 90-101.

Ukkusuri, S. V., Hasan, S., Luong, B., Doan, K., Zhan, X., Murray-Tuite, P., & Yin, W. (2016). A-RESCUE: An Agent based Regional Evacuation Simulator Coupled with User Enriched Behavior. *Networks and Spatial Economics*, 1-27.

Lim, M. B. B., Lim, H. R., Piantanakulchai, M., & Uy, F. A. (2016). A household-level flood evacuation decision model in Quezon City, Philippines. *Natural Hazards*, *80*(3), 1539-1561.

Song, Y., & Yan, X. (2016). A method for formulizing disaster evacuation demand curves based on SI model. *International journal of environmental research and public health*, *13*(10), 986.

Fu, H., & Wilmot, C. (2004). Sequential logit dynamic travel demand model for hurricane evacuation. *Transportation Research Record: Journal of the Transportation Research Board*, (1882), 19-26.

Cox, D. R. (1992). Regression models and life-tables. In *Breakthroughs in statistics* (pp. 527-541). Springer New York.

Kleinbaum, D. G. (1998). Survival Analysis, a Self-Learning Text. *Biometrical Journal*, *40*(1), 107-108.